Bone 141 (2020) 115575

Contents lists available at ScienceDirect

Bone

journal homepage: www.elsevier.com/locate/bone

Full Length Article

In vivo repeatability of homogenized finite element analysis based on multiple HR-pQCT sections for assessment of distal radius and tibia strength

Denis Schenk[a,*], Andrea Mathis[a], Kurt Lippuner[b], Philippe Zysset[a]

[a] ARTORG Center for Biomedical Engineering Research, University of Bern, Bern, Switzerland
[b] Department of Osteoporosis, Inselspital, Bern University Hospital, University of Bern, Bern, Switzerland

ARTICLE INFO

Keywords:
Bone strength
Distal radius
Distal tibia
Finite element analysis
HR-pQCT
Multiple sections
Osteoporosis
Repeatability

ABSTRACT

Introduction: Micro finite element analysis (μFE) is a widely applied tool in biomedical research for assessing *in vivo* mechanical properties of bone at measurement sites, including the ultra-distal radius and tibia. A finite element approach (hFE) based on homogenized constitutive models for trabecular bone offers an attractive alternative for clinical use, as it is computationally less expensive than traditional μFE. The respective patient-specific models for *in vivo* bone strength estimation are usually based on standard clinical high-resolution peripheral quantitative CT (HR-pQCT) measurements. They include a scan region of roughly 10 mm in height and are referred to as single-sections. It has been shown, that these small peripheral bone sections don't reliably cover the fracture line in Colles' fractures and therefore the weakest region at the radius. Recently introduced multiple section (multiple adjacent single-sections) measurements might improve the evaluation of bone strength, but little is known about the repeatability of hFE estimations in general, and especially for multiple section measurement protocols. Accordingly, the aim of the present work is to quantify repeatability of clinical *in vivo* bone strength measurement by hFE on multiple section HR-pQCT reconstructions at the distal radius and tibia.
Methods: Nineteen healthy Swiss women (43.6y ± 17.8y) and twenty men (48.2y ± 19.4y) were examined with HR-pQCT at 61 μm isotropic voxel resolution. Each subject was first scanned three times using a double-section (336 slices) at the distal radius and then three times using a triple-section (504 slices) at the distal tibia. The multiple section HR-pQCT reconstructions were graded for motion artefacts and non-linear hFE models (radius and tibia) and linear μFE models (only radius) were generated for estimation of stiffness and ultimate load. Then *in vivo* repeatability errors were computed in terms of root mean square coefficients of variation (CV).
Results: *In vivo* repeatability errors of non-linear hFE stiffness (S) and ultimate load (F) were significantly higher at the radius (S: 2.71% and F: 2.97%) compared to the tibia (S: 1.21%, F: 1.45%). Multiple section linear μFE at the radius resulted in substantially higher repeatability errors (S: 5.38% and F: 10.80%) compared to hFE.
Discussion/conclusion: Repeatability errors of hFE outcomes based on multiple section measurements at the distal radius and tibia were generally lower compared to respective reported single-section μFE repeatability errors. Therefore, hFE is an attractive alternative to today's gold standard of μFE models and should especially be encouraged when analyzing multiple section measurements.

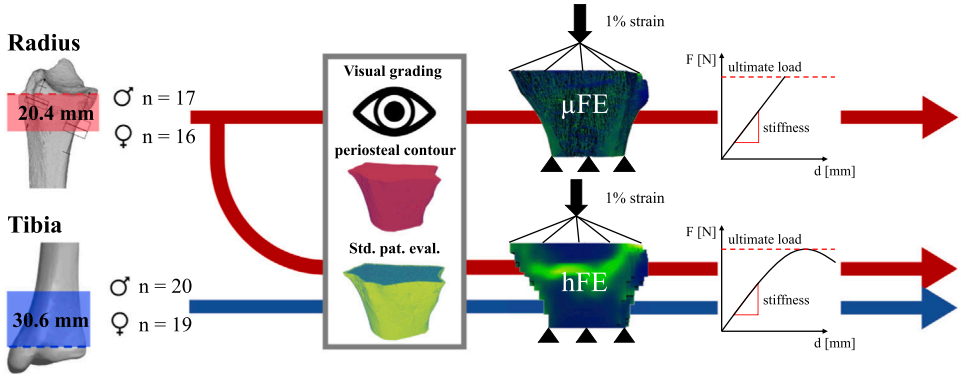
1. Introduction

Over the last few decades, micro finite element analysis (μFE) has matured into a widely applied tool for investigating the effects of specific diseases, treatments, activity or growth on the stiffness and strength of bones [1]. Bone strength measures estimated by μFE are as well gaining importance in diagnosis of osteoporosis and prediction of related fracture risks at anatomical locations assessed by high resolution peripheral quantitative computed tomography (HR-pQCT) [2]. *In vivo* μFE measures were able to differentiate between fracture and non-fracture groups in women and represent better surrogates of *in vivo* bone strength than bone mineral density (BMD) or architecture measures [4].

Varga and colleagues demonstrated, that FE models generated on a small peripheral bone section located in the fracture zone explain more variance in experimentally determined bone strength at the radius than respective, more complex, organ-level continuum FE models [5,6]. However, μFE models of such subsections demand a high amount of computational resources to solve for stiffness and even orders of magnitude higher resources to simulate the non-linear post-yield behavior

* Corresponding author.
E-mail address: denis.schenk@artorg.unibe.ch (D. Schenk).

https://doi.org/10.1016/j.bone.2020.115575
Received 21 February 2020; Received in revised form 5 August 2020; Accepted 5 August 2020
Available online 11 August 2020



	repeatability error [%]	
	radius	tibia
Tot.vBMD	2.11	0.65
hFE stiffness	2.71	1.21
hFE ult. load	2.97	1.45
μFE stiffness	5.38	–
μFE ult. load	10.80	–

Fig. 1. Graphical abstract showing from left to right: 1) measurement acquisition and schematic representation of the multiple section scan regions, 2) number of included samples differentiated by gender and measurement site, 3) visual grading and image processing according to the scanner manufacturer, 4) linear μFE and non-linear hFE simulations and 5) resulting repeatability errors as CV for radius and tibia measurements.

and compute ultimate load [7,8]. With *in vivo* resolutions of up to 60.7 μm, solving even linear single-section models of the distal radius can take up to several hours [9]. Thus, in most studies μFE models are limited to linear analyses. Instead of simulating the failure process of the respective bone segment, failure is defined to occur at an experimentally determined critical strain for a specified amount of bone tissue (*e.g.* Pistoia criterion [10]). An alternative finite element approach, using so called homogenized constitutive models for trabecular bone, has shown to accurately predict stiffness (R^2 = 0.85–0.86) and ultimate load (R^2 = 0.95–0.96) in comparison to *ex vivo* mechanical tests and μFE analysis [9,11]. These models are composed of elements about one order of magnitude larger than the underlying CT resolution and are computationally much more efficient than comparable μFE models, making them attractive for clinical use.

In contrast to helical or spiral CT scanners, HR-pQCT scanners are measuring bone tissue in so called sections. The width of such sections is determined by the width of the detector and its distance to the X-ray source. Today's standard clinical HR-pQCT measurement is a single section of approximately 1 cm starting at a fixed offset to defined anatomical landmarks [12,13]. In fact, this standardized clinical single-section at the distal radius does not reliably cover the entire fracture line in Colle's fractures [6,14]. From a mechanical standpoint, the scanned sections must include the weakest region of the measured bone to avoid overestimation of bone strength. Both, Varga et al. [6] and Mueller et al. [15] demonstrated that radius failure load decreases when the standard clinical section is shifted distally. Therefore, a new HR-PQCT measurement protocol was recently introduced [16] to scan a double-section at the radius (Fig. 2A). As the tibia is roughly 1.4 times longer than the radius [17], a triple-section is scanned at the distal tibia (Fig. 2B).

The validity of patient-specific μFE and homogenized FE (hFE) models has previously been investigated, showing high correlations with mechanical compression experiments of cadaveric samples [9–11,18,19]. In order to be of clinical use in detecting and monitoring changes in bone strength over time and during interventions or treatments, the measurement precision is fundamental [20]. The precision of a clinical setup depends on several factors and can be divided into three main areas: 1) errors and noise in the CT measurement procedure itself and variability in the subsequent image processing and FE pipelines; 2) errors introduced by *in vivo* measurements, including motion artefacts, miss-alignment of limb and reference line position and 3) BMD calibration errors in single center studies and cross-calibration errors in multicenter studies [21,22]. While calibration errors may degrade precision in the long term, motion artefacts and misalignment errors already influence short-term results (reproducibility) and even single measurements (repeatability). When evaluating follow-up measurements, misalignment errors can be reduced or even eliminated by image registration, while for single measurements, this is not possible. Clinical *in vivo* reproducibility of densitometric and morphological measures based on single-sections was found to be approximately 1% and 4.5%, respectively [12,23]. It is important to note that these results are based on images that were processed using the standard evaluation protocol of the manufacturer, which registers follow-up measurements using a 2D area matching method. The repeatability of single-section μFE results was investigated in several studies [22,23,25,36,37] with results summarized in [1]. Repeatability and reproducibility errors for single-center studies are less than 4.5% and 5.0% for radius stiffness and ultimate load and less than 3.7% and 4.28% for respective tibia outcomes. For multicenter studies, the errors for short- and long-term reproducibility can be increased by up to 1.8% and 1.4% for stiffness and strength [26]. In contrast, relatively little is known about the repeatability of hFE analyses. *Ex vivo* repeatability errors for hFE outcomes on cadaveric samples were evaluated by Hosseini and colleagues [11] to be about 1.7% and 2.25% for stiffness and ultimate load at the radius. These measurements were done using cadaveric forearms. They exclude motion artefacts, which are thought to be a significant error contributor.

Accordingly, the present study aims to quantify the single measurement repeatability of *in vivo* bone strength estimation by a newly available hFE pipeline on multiple section HR-pQCT reconstructions of the distal radius and tibia from a Caucasian population. We hypothesize that the single-measurement repeatability is identical at both measurement sites and independent of age.

2. Material and methods

The overall design of the study is summarized in Fig. 1.

2.1. Subjects

We recruited male and female participants aged 22–93 years. Subjects had no prior history of osteoporotic fracture, nor suffered from any medical condition or took any medication known to affect bone metabolism. The study protocol was approved by the Bernese ethics committee and all subjects gave written informed consent prior to participation. Twenty-seven degrees of freedom per group were considered to be sufficient to maintain the 95% upper confidence limit at 30% above the mean precision error in a subject group [20].

2.2. Scanning

All subjects were measured in a single visit at the University Department of osteoporosis in Bern, on a second-generation HR-pQCT scanner (XtremeCTII, Scanco Medical AG, Brüttisellen, Switzerland). Measurements were acquired at a nominal voxel size of 60.7 μm using standard *in vivo* scanner settings (60 kVp, 900 μA, 100 ms integration

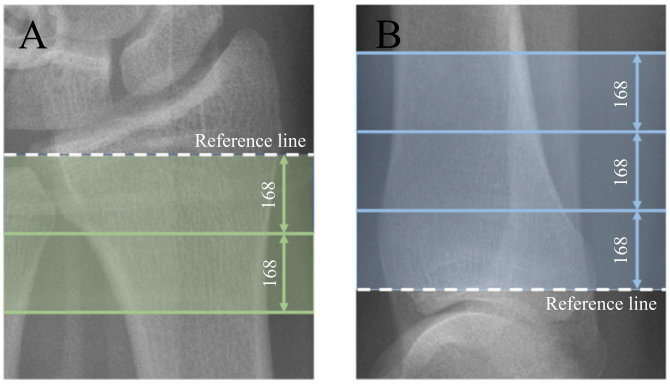


Fig. 2. Position of the reference line on the scout view images and visualization of multiple sections. (A) distal radius: the reference line was positioned at the proximal margin of the dense articular surface, formed with the scaphoid and lunate fossae of the radiocarpal joint. Then the adjacent proximal double-section (336 slices) was scanned without offset. (B) distal tibia: the reference line was positioned at the proximal margin of the dense structure formed by the tibial plafond and the adjacent proximal triple-section (504 slices) was scanned without offset. The size of the individual sections is only qualitatively indicated and doesn't reflect their real size.

Table 1
Abbreviations of HR-pQCT derived parameters for bone mass, geometry, microstructure and biomechanics.

Bone mass		
BMC	Bone mineral content	mg
Density		
Tot.vBMD	Total volumetric bone mineral density	mg/cm^3
Tb.vBMD	Trabecular volumetric bone mineral density	mg/cm^3
Ct.vBMD	Cortical volumetric bone mineral density	mg/cm^3
Tb.BV/TV	Trabecular bone volume fraction	%
Geometry		
Tb.Ar	Trabecular area	mm^2
Ct. Ar.	Cortical area	mm^2
Ct.Th	Cortical thickness	mm
Mask volume	Volume of the periosteal mask	cm^3
Microstructure		
Tb.N	Trabecular number	1/mm
Tb.Th	Trabecular thickness	mm
Tb.Sp	Trabecular separation	mm
Ct.Po	Cortical porosity	%
Biomechanics		
hFE S	Homogenized finite element stiffness	N/mm
hFE F	Homogenized finite element ultimate load	N
μFE S	Micro finite element stiffness	N/mm
μFE F	Micro finite element ultimate load	N

time).

First the subject's non-dominant forearm or leg of the same side (or the non-fractured side in case of a previous wrist or tibia fracture) was immobilized in a dedicated carbon fiber cast provided by the scanner manufacturer. The subject was positioned sideways to the scanner. The shoulder was laterally rotated to align the forearm with the scan axis. After the limb was positioned in the scanner, the operator took the standard scout view images and positioned the reference line at the proximal margin of the dense articular surface formed with the scaphoid and lunate fossae of the radiocarpal joint. Then the adjacent proximal double-section (336 slices) was scanned without offset (Fig. 2A). For distal tibia measurements, the subject was positioned along the scan axis of the scanner. The hip was flexed, and the knee was slightly bent. After taking the standard scout view, the reference line was positioned at the proximal margin of the dense structure formed by the tibial plafond and the adjacent proximal triple-section (504 slices) was scanned without offset (Fig. 2B). For each subject, we first performed three measurements at the radius, followed by three measurements at the tibia. After each scan, the carbon cast was removed and reapplied and the subject was repositioned. The six measurements of each individual subject were performed by the same operator and there were two operators in total.

2.3. Grading

After reconstruction, one single operator graded all measurements for their motion artefacts on a scale of 1 (G1: no motion artefacts) to 5 (G5: extreme motion artefacts) [27]. For double-section measurements the first and last slice (1 and 336) were graded individually, while for triple-section measurements, the first, middle and last slice (1, 252 and 504) were graded. The final grade of each measurement was defined as the highest individual slice grade.

2.4. Image processing

All measurements, independent of their quality grading, were processed using the standard clinical workflow for XtremeCTII images implemented on the scanner software (IPL Scanco Module 64-bit Version V5.16/FE-v02.02), provided by the scanner manufacturer. First, an automatic contouring algorithm was applied to define the periosteal contour. Erroneous contours including sections of the ulna or fibula were excluded from further evaluations. The remaining measurements were evaluated with the standard patient evaluation protocol for identifying the endosteal surface and segment the cortical and trabecular regions. This included Gauss filtering operation (sigma = 0.8, support = 1 voxel) and thresholding (cortical bone: 450 mgHA/cm^3, trabecular bone: 320 mgHA/cm^3). The standard output variables (abbrev. acc. to [28], summarized in Table 1) were classified into density measures (Tot.vBMD, Tb.vBMD, Ct.vBMD and Tb.BV/TV), geometry variables (Tb.Ar, Ct.Ar and Ct.Th) and microstructural parameters (Tb.N, Tb.Th, Tb.S and Ct.Po). Additionally, a custom-made IPL script computed the volume of the periosteal contour (Mask volume), as well as bone mineral content (BMC) as the product of Tot.vBMD and Mask volume. All parameters were evaluated *without* the usually performed cross-sectional area registration (CSA) for follow-up measurements.

2.5. Homogenized finite element analysis

Bone stiffness and ultimate load were estimated using the validated and standardized non-linear homogenized finite element method published by Arias-Moreno et al. [9] and Hosseini et al. [11]. A short overview is provided hereafter. Periosteal contours were downscaled with a factor of 28 to 8-node brick elements with 1.7 mm edge length. For each element, bone material properties (Young's modulus ε_0 = 19.01 GPa, shear modulus μ_0 = 7.851 GPa, compressive strength σ_0^- = 166 MPa, maximum tensile strength σ_0^+ = 131 GPa and maximum shear strength τ_0 = 67.3 GPa) were homogenized based on vBMD and transverse isotropy using the mean intercept length (MIL [29]) fabric tensor. For elements containing only cortical bone, the MIL tensor was set to identity (isotropy). The constitutive law included linear elasticity followed by yielding and the simultaneous accumulation of damage and irreversible strains. A detailed description is provided in [30]. Displacements in all 3 DOF of the most proximal element nodes and the two in-plane displacements of the most distal element nodes were fixed to zero. The models were loaded with a uniformly axial displacement on the most distal element nodes up to 1% of strain. Reaction force and displacement of the top surface were recorded. Stiffness was computed as the initial slope of the resulting force-displacement curve and ultimate load was defined to be the maximal recorded reaction force. All models were built and evaluated on the XtremeCTII scanner using 1 core of an HP Integrity Server rx2800 i4.

2.6. Micro finite element analysis

All included radius measurements were also analyzed for stiffness and ultimate load using the μFE pipeline implemented on the scanner software [9]. The distal tibia triple-sections were not evaluated due to the high computational effort for solving them. In brief, the μFE models were generated by a direct voxel conversion approach based on the segmentation of the standard evaluation. Linear elastic material properties (Young's modulus E = 10 GPa, Poisson's ratio ν = 0.3) were assigned to all elements. The most proximal element nodes were fixed in all 6 DOF. In-plane transversal displacements of the most distal element nodes were restricted and a compressive strain of 1% was applied. Stiffness was then computed as the total reaction force divided by the total displacement. Ultimate load or failure load was defined according to the Pistoia criterion to occur when a minimum of 2% of the tissue volume is loaded beyond a critical strain of 0.7%. Although Arias-Moreno et al. adapted this failure criterion for double-sections, we decided to use the standard values.

2.7. Statistical evaluations

Differences in age distribution between male and female samples were tested with non-parametric two-sided Mann-Whitney-Wilcoxon tests after confirming non-normality of variables with Shapiro-Wilk tests. Then we computed linear regressions from the means of repeated measurements of hFE ultimate load with hFE stiffness, density parameters and bone mass parameters.

The further statistical analyses are divided into three distinct parts: Primary analysis, secondary analyses, and an explorative analysis with several statistical analyses which were underpowered.

2.7.1. Primary analysis

The *in vivo* single measurement repeatability errors of the described techniques and of individual subjects were computed based on [20] as root mean square coefficients of variation (CV). Subjects had different degrees of freedom, resulting from the exclusion of some individual measurements for which the automatic periosteal contouring procedure failed. Therefore, the mean of individual repeatability errors doesn't correspond to the CV of the technique.

2.7.2. Secondary analyses

Differences in grading score of both measurement sites were tested with an unpaired Mann-Whitney test on the maximum grading score of repeated measurements, differentiated by the measurement site (radius/tibia).

To test influences of age and measurement site on individual repeatability errors, we built a multiple linear regression model with the log transformed individual repeatability errors of all outcome variables as response and age ($\leq$50y/ > 50y), measurement site (radius/tibia) and their interaction as predictors. The normality assumption for the response was tested with Q-Q plots and Shapiro Wilk tests.

$$\log(CV) = \beta_0 + \beta_1 * site + \beta_2 * age + \beta_3 * (age * site)$$

To compare individual repeatability errors between hFE and μFE outcomes (stiffness and ultimate load), we computed the differences in individual repeatability errors between the respective sample pairs. The distribution of the differences was significantly different from normal distribution (Shapiro test), so we bootstrapped the mean difference with 10,000 replicates. The bootstrapped means were normally distributed. Accordingly, we computed the respective 95% confidence intervals (CI) assuming normal distribution. These evaluations were performed on the mixed population (female + male) and only for the radius.

The same bootstrapping procedure was applied to compare individual repeatability errors between hFE outcomes (stiffness and ultimate load) and Tot.vBMD. Again, these evaluations were performed on the mixed population, but in contrast, we differentiated the CIs by measurement site (radius/tibia).

2.7.3. Explorative analyses

To test a potential influences of grading scores on individual repeatability errors of hFE outcomes (radius + tibia), we performed a Kruskal-Wallis test with groups of maximum grading score of the repeated measurements (G1 – G5). In case this test was significant, we performed pairwise comparisons between maximum grading levels with post-hoc Wilcoxon rank sum tests with *p*-value correction according to Holm.

The statistical analyses were performed in R (The R foundation for statistical computing, Vienna, Austria, version 3.6.3) and $p < .05$ was considered to be statistically significant.

3. Results

3.1. General

A total of 19 women (43.6y ± 17.8y) and 20 men (48.2y ± 19.4y) with age range from 22 to 93 (45.9y ± 18.6y) were scanned each three times at the distal radius and tibia, except for two participants. One woman was only measured twice at the radius and at the tibia, and one man was only measured three times at the radius and twice at the tibia. We had to exclude a total of 23 measurements (men radius: 10/60, men tibia: 2/59, women radius: 11/56, women tibia: 0/56) because the automatic periosteal masking included parts of the ulna or fibula making them unusable for further evaluation, without any manual correction. The included measurements resulted in a total of 33 and 37 DOF for men radius and tibia, and 29 and 37 DOF for women radius and tibia respectively. The characteristics of the study population for radius and tibia is summarized in the attachment (Table 5 and 6).

The age in men and women was not significantly different ($p = .21$). Thus, no age groups were differentiated when testing for differences in individual repeatability errors.

The distribution of quality grades of all included measurements is shown in Fig. 3. None (0/7) of the measurements with quality grade G5 and roughly 47% (7/15) of the measurements with quality grade G4 resulted in appropriate periosteal contours. About 4% and 2% of men radius and tibia, and 7% and 4% of all women radius and tibia scans were graded with G4.

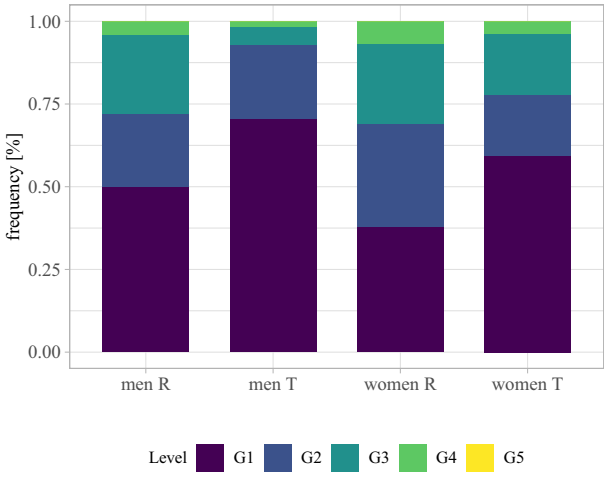


Fig. 3. Distribution quality grades of all *included* measurements, differentiated by gender (women/men) and measurement site (R = radius, T = tibia). None of the measurements with quality grading G5 was included in the study, therefore the frequencies of G5 in this representation are zero.

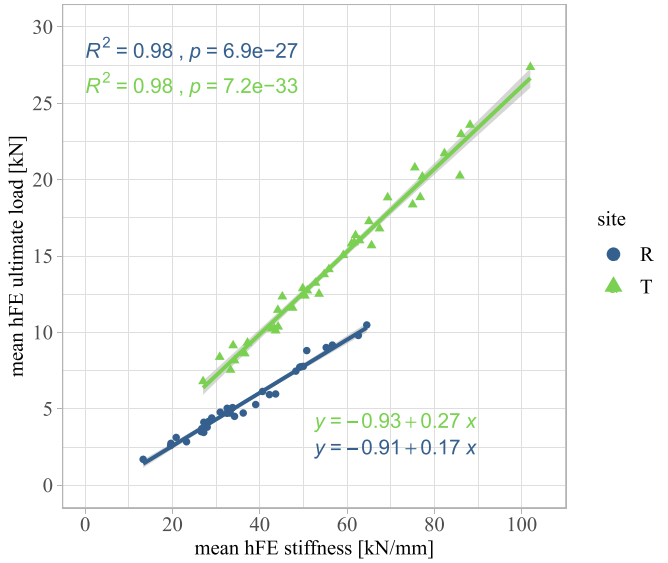


Fig. 4. Linear regression of hFE stiffness with hFE ultimate load at the radius (blue) and the tibia (green). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.2. Linear regressions with hFE ultimate load

Despite the actual simulation of bone failure, linear correlation of hFE stiffness with ultimate load resulted in the highest coefficients of determination at both the radius (R^2 = 0.98, intercept = −908, slope = 0.174) and the tibia (R^2 = 0.98, intercept = −930, slope = 0.27) (Fig. 4). The regressions have a different slope and do not intersect each other. Among all subject means of density and bone mass parameters, linear regression of hFE ultimate load was highest with BMC at the radius (R^2 = 0.82, intercept = −2177, slope = 3.45) and with Tb.BVTV at the tibia (R^2 = 0.88, intercept = −12,578, slope = 87,926). Linear regression of hFE ultimate load at the distal radius with hFE ultimate load at the distal tibia resulted in a lower coefficient of determination (R^2 = 0.76, intercept = 4231, slope = 1.94).

3.3. Primary analysis

Short-term *in vivo* repeatability errors are summarized in Table 2 for the radius and tibia, respectively. Repeatability errors of density outcomes ranged from 1.02–2.11% at the radius and 0.65–2.97% at the tibia. In comparison, the repeatability of hFE stiffness and ultimate load was 2.71% and 2.97% at the radius, and 1.21% and 1.44% at the tibia, respectively. Repeatability errors of μFE stiffness and ultimate load at the radius were 5.38% and 10.80%, respectively.

3.4. Secondary analysis

The maximum grading score of repeated measurements was significantly lower for tibia measurements compared the radius (p = .002).

The distribution of log transformed individual repeatability errors was generally not significantly different from normality distribution, except for Ct.vBMD, Tb.N and Mask Volume. Stiffness and ultimate load estimated by hFE at the tibia generally resulted in lower repeatability errors compared to the same variables at the radius. The computed linear regression models revealed this trend to be significant for hFE stiffness (p < .001) and ultimate load (p = .014). Among all other outcome variables, linear regression models revealed significant differences in log transformed individual repeatability errors between the radius and tibia for BMC (p = .019), Tot.vBMD (p < .001), Tb.Ar (p < .001), Mask Volume (p = .005) and Ct.Po (p < .001).

Individual repeatability errors of hFE stiffness and ultimate load at the radius and tibia and of μFE stiffness and strength at the radius are shown in Fig. 5. The 95% CI of the bootstrapped paired differences in individual repeatability errors between hFE and μFE at the radius were [1.58, 4.66] for stiffness and [2.70, 7.72] for ultimate load, respectively. Zero was not included in either of the two CIs.

Finally, comparison of individual repeatability errors of Tot.vBMD and hFE outcomes from the mixed population is shown in Fig. 6 for radius and tibia, respectively. We found the 95% CI of the bootstrapped paired differences in individual repeatability errors at the radius as follows: [0.21, 1.03] between hFE stiffness and Tot.vBMD, [0.04, 1.06] between hFE ultimate load and Tot.vBMD and [−0.47, 0.33] between hFE ultimate load and stiffness. The respective 95% CI at the tibia were: [0.46, 0.85] between hFE stiffness and Tot.vBMD, [0.61, 1.01] between hFE ultimate load and Tot.vBMD and [0.03, 0.28] between hFE ultimate load and stiffness.

3.5. Explorative analysis

Individual repeatability errors of hFE stiffness and ultimate load categorized by maximum grading score of repeated measurements are shown in Fig. 7. The Kruskal test was only significant for hFE stiffness (p = .007). The pairwise comparison between the maximum grading groups revealed a significant difference in individual repeatability

Table 2
Repeatability errors (CV) for radius and tibia measurements and corresponding p-values of the multiple linear regression models of individual repeatability errors for predictor "measurement site" (boldface: significant at 95% level).

		Radius	Tibia	
Number of subjects[a]		33	39	
DOF		62	74	
		CV(%)	**CV(%)**	**p-Value** Predictor: site
Bone mass				
BMC	mgHA	0.896	1.324	**0.019**
Density				
Tot.vBMD	mg HA/cm^3	2.109	0.652	**1.50e-07**
Tb.vBMD	mg HA/cm^3	1.025	2.969	0.951
Ct.vBMD	mg HA/cm^3	1.428	0.945	0.490[b]
Tb.BV/TV	–	1.445	2.967	0.194
Geometry				
Tb.Ar	mm^2	3.150	5.489	**7.7e-04**
Ct.Ar	mm^2	1.807	3.918	0.445
Ct.Th	mm	2.695	3.044	0.411
Mask volume	mm^3	2.429	1.238	**5.1e-03**[b]
Microstructure				
Tb.N	1/mm	2.448	2.277	0.406[b]
Tb.Th	mm	1.022	1.997	0.929
Tb.Sp	mm	1.552	1.985	0.113
Ct.Po	%	13.099	7.394	**1.40e-04**
Biomechanics				
hFE stiffness	N/mm	2.708	1.205	**6.73e-03**
hFE ultimate load	N	2.967	1.448	**0.014**
μFE stiffness	N/mm	5.380	–	–
μFE ultimate load	N	10.803	–	–

Values are expressed as root mean square coefficient of variations CV (in %); Abbreviations acc. to [28].

[a] Number of subjects: all subjects with either DOF 1 or 2 (3 measurements = 2 DOF, 2 measurements = 1 DOF).

[b] For these variables, the performed Shapiro test revealed a distribution significantly different from normal distribution. This undermines the normality assumption of the response in linear regression models.

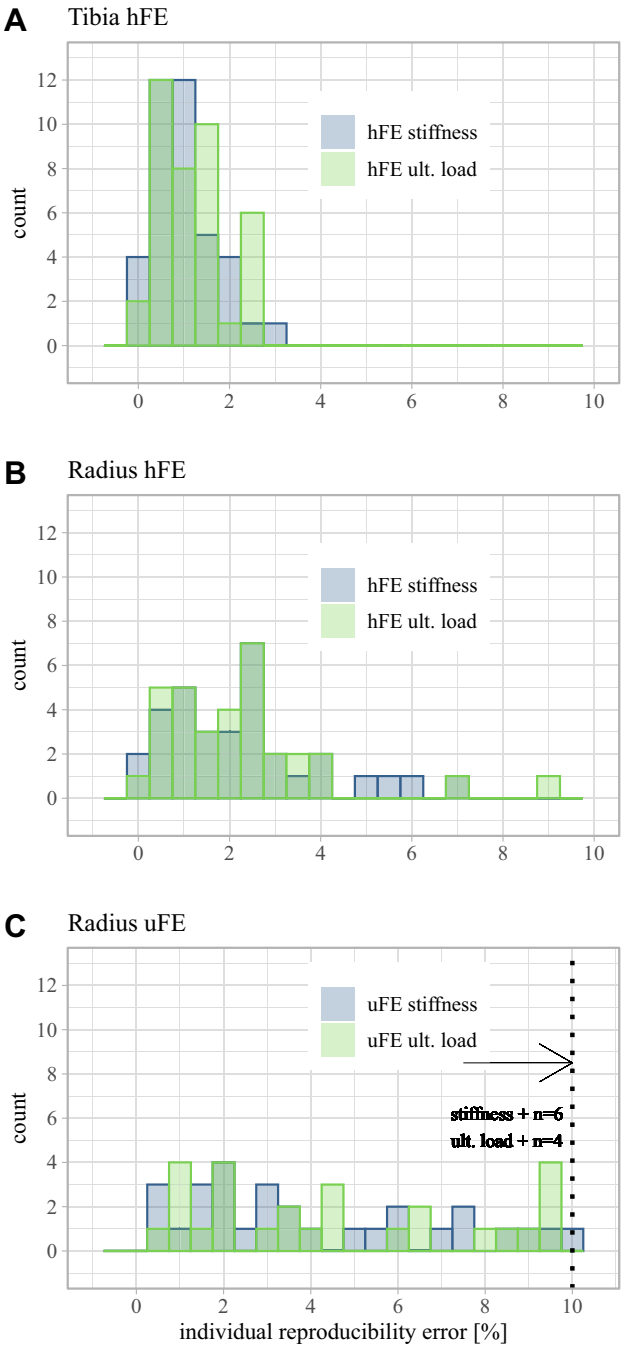


Fig. 5. Histograms of individual repeatability errors of hFE stiffness (blue) and ultimate load (green) (A) at the tibia and (B) the radius and (C) of μFE stiffness and ultimate load at the radius. Panel C was cut at x = 10.0 to allow easier comparison with A and B. Missing counts are listed above the arrow for μFE stiffness and ultimate load. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

errors between group 1 and 3. These tests were explorative, and some groups didn't contain enough samples to be sufficiently powered.

4. Discussion

4.1. General

This study evaluated the single-measurement *in vivo* repeatability errors of a novel and standardized hFE methodology based on second-generation HR-pQCT reconstructions of multiple sections of the distal radius and tibia in the healthy Caucasian population. A summary of repeatability errors of previous studies is provided in Table 3 for the radius and Table 4 for the tibia. To our best knowledge, we report here the first single-measurement *in vivo* repeatability errors for multiple section measurements and for the hFE method. All presented parameters were evaluated without the use of the 2D cross-sectional area registration provided by the scanner manufacturer. Among the summarized studies, only Ellouz and colleagues [22] forego any registration as well. Additionally, only Chiba and colleagues [31] reported repeatability errors based on second-generation HR-pQCT reconstructions.

As reported in previous studies on single-section measurements, our multiple section repeatability errors for density-based measures tend to be lower than for geometry and microstructural measures, independent on measurement site [32].

Although the grading method proposed by Pialat and colleagues [27] recommends to only include measurements up with gradings G1–G3, we decided prior to the study to only include measurements where we could mask the periosteal contour automatically, without manual correction and independent on their grading score. This offers a simple exclusion criterion and an efficient workflow for studies with larger sample size. The alternative semi-automatic masking procedure provided by the scanner manufacturer can be very time consuming and increases the influence of operators on outcomes. However, it remains unclear, if this procedure could even have been applied to the excluded images, as on most of them, motion artefacts were severe (G4 and G5). From a total of 116 radius and 115 tibia measurements, we had to exclude roughly 18% at the radius, but less than 2% at the tibia. Acquiring high quality images and minimizing motion artefacts reflected a practical issue that needs to be incorporated into the study design. Significantly lower grading scores at the tibia compared to the radius indicate that tibia measurements were less prone to subject motion. We think that this can mostly be attributed to the more stable and comfortable positioning of the patient during measurements. Accordingly, the high exclusion rate at the radius was most probably not much influenced by the chosen exclusion criterion.

4.2. Linear regression of hFE stiffness and hFE ultimate load

The linear regressions of hFE ultimate load with hFE stiffness resulted in very high coefficients of variation at both the radius and tibia. This basically offers the possibility to estimate ultimate load based on stiffness and reduces the processing time substantially. However, the relation between ultimate load and stiffness is very much depending on the underlying structure and material. As soon as these factors change, the relation changes as well. This can clearly be seen on Fig. 4, where the respective regressions of radius and tibia measurements differ substantially in slope. Moreover, such general models could not be applied to any individual cases including pathologies, such as lytic tumors. Therefore, a direct simulation of bone strength seems to be preferable in clinical applications, handling individual pathological cases.

4.3. Repeatability compared to literature

The only reported *in vivo* study using single-section measurements without any registration was performed on a first-generation HR-pQCT scanner [22]. In comparison, the repeatability errors on second-generation HR-pQCT measurements at the radius and tibia are about 20% and 50% lower for Tot.vBMD, 30% lower and 65% higher for Tb.vBMD and 80% higher and 15% lower for Ct.vBMD, respectively. Compared to previous studies reporting CSA-based registered first-generation HR-pQCT measures, our repeatability errors for density-based measures are in general similar or slightly higher at both the radius and the tibia, except for Tb.vBMD. The present repeatability errors of Tb.vBMD at the tibia were consistently higher compared to all other summarized studies. In multiple sections, the gradient of Tb.vBMD from the proximal to the distal end is higher, compared to single-sections. Accordingly, a

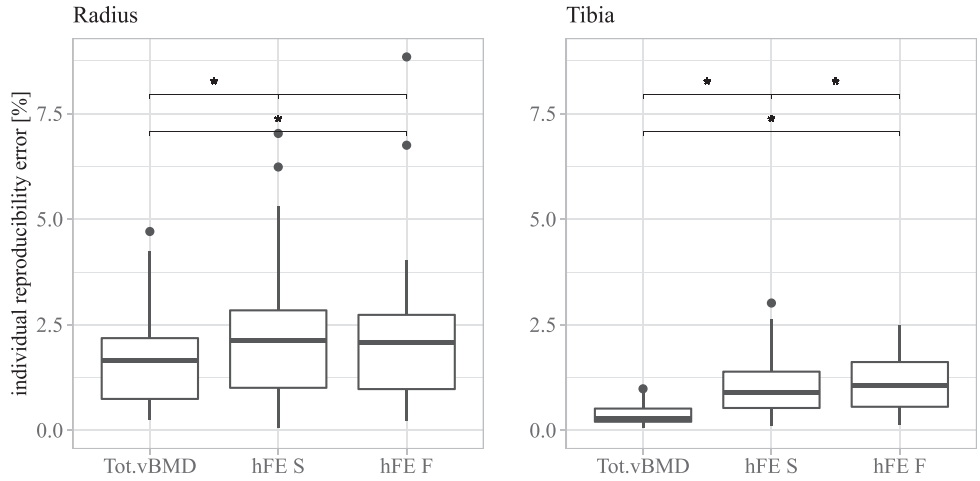


Fig. 6. Comparison of individual repeatability errors of hFE outcomes with Tot.vBMD for radius and tibia on the mixed population (men and women). (*) indicates, that zero was not included in the computed CI of the respective paired differences.

small shift in the scanned region will induce higher differences between measurements and result in higher repeatability errors.

Our repeatability errors of microstructural measures are lower compared to CSA-registered repeatability on first-generation HR-pQCT scanners and about similar compared to CSA-registered repeatability on second-generation HR-pQCT. Improvements in spatial resolution and measurement algorithms in the second-generation HR-pQCT lead to better or equal repeatability for most outcomes compared to the first generation, especially for trabecular microstructure [33]. Further, CSA-registration does not significantly improve microstructural measures [22], explaining why our microstructural repeatability compares well with Chiba and colleagues, who used CSA registration.

4.4. Age

We found that in our population, repeatability errors of hFE stiffness and ultimate load were not significantly different between the two age groups. This is in agreement with Paggiosi et al. [25], who investigated μFE repeatability at the distal radius and tibia for three different age groups. They only found differences in the repeatability of estimated failure loads between the oldest (> 70y) and youngest (16–18y) group in men.

4.5. Grading

The repeatability errors of the maximum grading groups were only significantly different for hFE stiffness at the tibia. Although the respective statistical test was underpowered, at the tibia better grading scores seem to lead to lower repeatability errors. We could not detect this trend at the radius. The larger volume of the tibia might magnify motion artefacts visible by the eye. Consequently, the visual grading at the tibia might be more sensitive, which leads to a more realistic representation of the influence of grading on repeatability errors. We believe that grading is in fact influencing repeatability errors. Apart from the small sample size, the applied grading method might not be objective and sensitive enough to accurately describe this relation. Quantitative and more objective grading methods, such as the one proposed by Sode and colleagues [34], might improve image quality metrics for HR-pQCT measurements.

4.6. In vivo versus ex vivo repeatability errors for hFE

Hosseini and colleagues were the only ones reporting any repeatability errors for hFE stiffness and ultimate load at the radius [11]. They scanned double-sections of cadaveric forearms on second-generation HR-pQCT scanners and repositioned the cadavers between the scans. They reported cadaveric repeatability errors roughly 1% lower than the

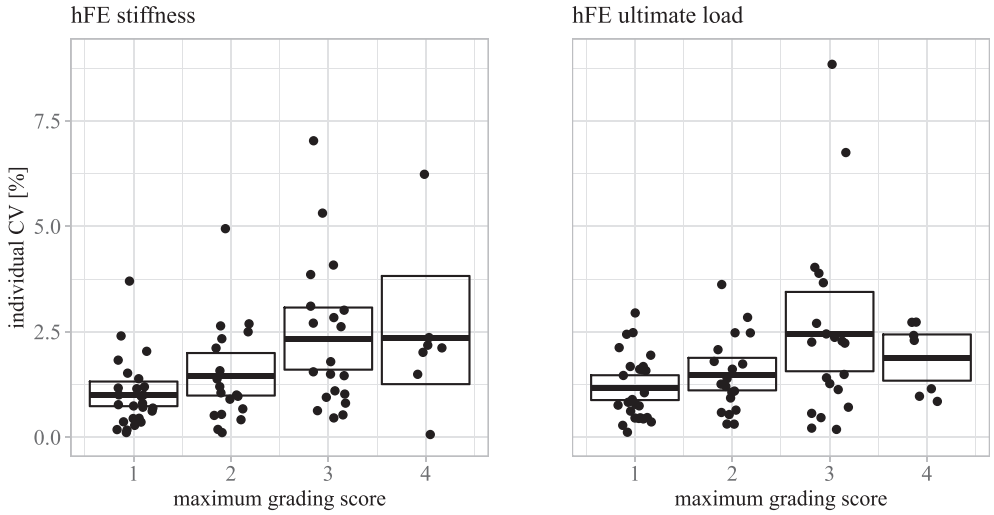


Fig. 7. Individual repeatability errors of hFE stiffness (left) and hFE ultimate load (right) categorized by maximum grading score of repeated measurements.

Table 3
Summary of distal radius repeatability errors (CV) of the present and previous studies. If gender or age groups were differentiated, repeatability errors are summarized as a range. Abbreviations: BL = baseline, Std. single sec. = Standard single-section.

	Authors	Present study	Ellouz et al. [22]	Kawalilak et al. [36]	Bonaretti et al. [37]	Boutroy et al. [12]	MacNeil & Boyd [23]	Paggiosi et al. [25]	Kawalilak et al. [38]	Chiba et al. [31]
	Year	2020	2014	2016	2017	2005	2008	2014	2014	2018
	CT generation	Second	First	First	First	First	First	First	First	Second
	Subjects	33 men and women	15 healthy subjects (21–47y)	34 postmen. Women (74 + −7y)	57 men and women	15 healthy women (aged 21–47 y)	15 men (20–37y) and 15 women (20–40y)	180 men and women (3 age groups)	29 young healthy adults (27 + −9y)	15 healthy men and women (20–74y)
	Scans	Triple BL scans with repositioning	3 separate scans within 1 mo	2 separate scans at BL and 10 d	Double BL scans with repositioning	3 separate scans within 1 mo	3 separate scans at BL, 1 wk. and 4 mo	Double BL scans with repositioning	2 separate scans at BL and 10 d	3 separate scans at BL, 1 w and 4 wk
	Scan region	Double-section	Std. single-sec.	Std. single-sec.	Std. single-sec.	Std. single-sec.	Std. single-sec.	Std. single-sec.	Std. single-sec.	Std. single-sec.
	CSA reg.	No	No	No	Yes	Yes	Yes	Not specified	Yes	Yes
Density	Tot.vBMD	2.11	2.6	–	3.3	0.9	0.39–0.62	0.7–1.1	1.6	1.02
	Tb.vBMD	1.03	1.5	–	1.63	1.0	0.49–0.66	0.9–1.4	0.9	1.54
	Ct.vBMD	1.43	0.8	–	2.35	0.7	0.34–0.40	2.3–5.5	1.2	0.56
Geometry	Tb.Ar	3.15	–	–	–	–	–	–	0.6	0.38
	Ct.Ar	1.81	–	–	–	–	–	–	3.1	1.27
	Ct.Th	2.70	3.7	–	7.88	1.2	0.66–1.35	2.0–3.9	3.1	1.17
Microstructure	Tb.N	2.45	2.8	–	4.86	3.0	3.36–3.60	4.3–6.9	5.7	2.04
	Tb.Th	1.02	3.3	–	–	3.2	3.10–3.41	3.8–6.6	5.9	1.00
	Tb.Sp	1.55	2.8	–	–	2.8	3.28–4.40	4.4–7.0	5.7	1.63
	Ct.Po	13.10	11.8	–	12.02	–	–	10.5–20.3		13.28
Biomechanics	μFE stiffness	5.38	3.6	3.3–4.4	–	–	1.44–2.36	3.5–4.5	–	–
	μFE ult. load	10.80	–	2.8–5.0	2.88	–	1.44–2.47	3.3–3.7	–	–
	hFE stiffness	2.71	–	–	–	–	–	–	–	–
	hFE ult. load	2.97	–	–	–	–	–	–	–	–

Table 4
Summary of distal tibia repeatability errors (CV) of the present and previous studies. If gender or age groups were differentiated, repeatability errors are summarized as a range. Abbreviations: BL = baseline, Std. single sec. = Standard single-section.

	Author	Present study	Ellouz et al. [22]	Kawalilak et al. [36]	Bonaretti et al. [37]	Boutroy et al. [12]	MacNeil et al. [23]	Paggiosi et al. [25]	Kawalilak et al. [38]	Chiba et al. [31]
	Year	2020	2014	2016	2017	2005	2008	2014	2014	2018
	CT generation	Second	First	First	First	First	First	First	First	Second
	Subject	39 men and women	15 healthy subjects (21–47y)	34 postmen. Women (74 + −7y)	63 men and women	15 healthy women (21–47 y)	15 men (20–37y) and 15 women (20–40y)	180 men and women (3 age groups)	29 young healthy adults (27 + −9y)	15 healthy men and women (20–74y)
	Scan	Triple BL scans with repositioning	3 separate scans within 1 mo	2 separate scans at baseline and 10 d	Double BL scans with repositioning	3 separate scans within 1 mo	3 separate scans at BL, 1 wk. and 4 mo	Double BL scans with repositioning	2 separate scans at BL and 10 d	3 separate scans at BL, 1 wk. and 4 wk
	Scan region	Triple-section	Std. single-sec.	Std. single-sec.	Std. single-sec.	Std. single-sec.	Std. single-sec.	Std. single-sec.	Std. single-sec.	Std. single-sec.
	CSA reg.	No	No	No	Yes	Yes	Yes	Not specified	Yes	Yes
Density	Tot.vBMD	0.65	1.5	–	0.90	1.3	0.48–1.03	0.2–0.6	0.7	1.39
	Tb.vBMD	2.97	1.8	–	0.93	1.5	0.54–0.91	0.2–0.6	0.8	1.30
	Ct.vBMD	0.95	1.1	–	1.54	0.9	0.25–0.58	0.1–1.7	0.4	0.68
Geometry	Tb.Ar	5.49	–	–	–	–	–	–	–	0.61
	Ct.Ar	3.92	–	–	–	–	–	–	–	2.36
	Ct.Th	3.04	1.4		2.52	0.9	0.86–2.24	0.4–1.5	0.9	1.11
Microstructure	Tb.N	2.28	4.4	–	5.71	3.8	3.59–3.89	3.0–5.4	4.0	2.41
	Tb.Th	2.00	4.8	–	–	4.4	3.62–3.85	2.9–5.3	3.8	0.75
	Tb.Sp	1.98	4.2	–	5.61	4.3	3.51–4.00	3.0–5.4	4.0	1.70
	Ct.Po	7.39	8.0	–	5.61	–	–	–	–	11.03
Biomechanics	μFE stiffness	–	2.3	2.1–3.7	–	–	1.6–3.55	1.0–3.0	–	–
	μFE ult. load	–	–	2.5–2.9	2.76	–	1.86–4.28	0.6–2.3	–	–
	hFE stiffness	1.20	–	–	–	–	–	–	–	–
	hFE ult. load	1.45	–	–	–	–	–	–	–	–

ones we found. This trend is expected as they used about 40% smaller FE elements and cadaveric measurements exclude motion artefacts normally occurring during *in vivo* procedures.

4.7. Double-section μFE versus double-section hFE and single-section μFE

The presented *in vivo* repeatability errors for double-section μFE ultimate load at the radius are about 1.5 times higher than what Paggiosi and Kawalilak reported on *in vivo* single-sections and about 2–3 times higher than the corresponding hFE repeatability errors. In multiple section models, there exists a perturbation in the bone structure at the boundary of two adjacent sections, due to subject motion. This leads to discontinuities in the subsequent μFE model, causing regions with high strain at the interface of the two sections. In our μFE simulations, we estimated ultimate load using the well-known criteria of Pistoia and colleagues [10], which is based on strain measures. The strain regions near the interface are not representing the *in vivo* biomechanical situation and are therefore artificial. They undermine any failure criterion based on strain measures, leading to high repeatability errors. Our results suggest that the nature of hFE models can compensate for discontinuities at the intersection of multiple stack models, as they are mainly based on homogenized vBMD and fabric tensors rather than their segmentation and consist of elements an order of magnitude larger than the underlying CT resolution.

4.8. Multiple section hFE versus single-section μFE

Compared to today's gold standard of single-section μFE for *in vivo* estimation of bone strength, repeatability errors of the present multiple section hFE outcomes were generally lower at both measurement sites, except for some specific age-groups in Paggiosi et al. at the tibia and the μFE repeatability errors reported by MacNeil and colleagues at the radius. Followingly, bone strength estimated by multiple section hFE might be an interesting alternative to μFE models in terms of clinical repeatability and processing time. FE analysis on multiple sections is less dependent on the applied boundary conditions, and the region of interest for follow up measurements is much larger. Multiple sections also offer the possibility to define patient size-dependent regions of interest during post-processing instead of during a clinical examination.

4.9. hFE radius versus tibia

Individual repeatability errors of hFE outcomes were lower at the tibia compared to the radius. This might indicate the use of tibia measurements as surrogate for radius measurements. However, such relations are not convincing as suggested by the rather low coefficient of determination between hFE outcomes at the radius and tibia. Measurements at one site do not necessarily reflect measurements at the other. Errors introduced by doing so would be far greater than improvements in repeatability.

4.10. hFE versus Tot.vBMD

Individual repeatability errors of vBMD measures were significantly lower than for corresponding hFE stiffness and ultimate load repeatability errors at both the radius and tibia. This is not surprising, as the homogenized trabecular bone material properties used for hFE analysis depend on power functions of local vBMD with an exponent larger than 1.

4.11. Clinical repeatability

Major contributors to *in vivo* measurement precision and repeatability have been previously reported by MacNeil and Boyd [23]. They include subject positioning and repositioning, motion artefacts, measurement VOI identification, and inherent variability or machine precision. For follow-up measurements it is essential to mention that hFE outcomes are sensitive to changes in BMD calibration. Such changes may occur slowly over time when the calibration equation is not regularly updated based on quality control measurements, but as well from smaller daily variabilities. Unpublished results of ours suggest that a change in vBMD of 1% (maximum permissible calibration drift according to the scanner manufacturer) may change hFE stiffness and ultimate load by up to 2%. These errors are in the range of our reported hFE repeatability errors and emphasize the major importance of BMD calibration and daily quality control measurements. While hFE outcomes can retrospectively be corrected for long-term BMD calibration drifts, the daily variability (in our case: SD of residuals = 0.63 mgHA/ccm) will remain and add up to short term *in vivo* repeatability errors.

Finally, Caksa and colleagues recently showed that it is crucial to understand the potential impact of overlaying soft tissue on the accuracy of HR-pQCT measurements. In an experimental setup they found a decrease in measured Tot.vBMD of up to 3.7% when the overlaying circumferential fat on a phantom was increased from 0 mm to 12 mm [35]. This would lead to changes in hFE stiffness and ultimate load of more than 7%, a multiple of the estimated hFE *in vivo* repeatability errors themselves.

4.12. Limitations

Our study has several limitations. In order to investigate single measurement repeatability errors, we acquired up to three HR-pQCT measurements at the distal radius and tibia back-to-back. We compared these errors to several reference studies from literature. Among these studies, only two acquired their HR-pQCT measurements back-to-back. The others acquired the respective measurements over a time period of up to four months. Therefore, we decided to categorize the present repeatability errors as single measurement errors, rather than short-term errors. Acquiring measurements back-to-back excludes any errors resulting from a potential calibration drift. Such drifts might lead to higher repeatability errors when measurements are acquired over a longer time period.

Another limitation is the exclusion of patients with a previous diagnosis of osteoporosis or any other bone disease. Osteoporotic patients may have higher repeatability errors. We expect these differences to be higher for geometry and microstructural parameters, compared to density-based measures, as the former rely on a segmentation of vBMD measurements.

We could not find any significant differences in repeatability errors between our two age groups. This does not mean that there are no differences but may reflect that they were too small to be detected with our sample size. Additionally, a larger sample size would have allowed differentiating more age categories.

Lastly, we had to perform a BMD recalibration during the measurement period, because the daily quality control measurements were out of bound. This can potentially affect the evaluated parameters, but not the assessed repeatability errors, as none of the subjects contained pre- and post-recalibration measurements.

5. Conclusion

In conclusion, we have established the single measurement *in vivo* repeatability of second-generation HR-pQCT based hFE stiffness and ultimate load of multiple sections at peripheral skeletal sites. We found that in the selected population hFE repeatability errors are not significantly different between the two age groups and are significantly lower at the tibia compared to the radius. Despite the lower repeatability errors, measurements of ultimate load at the tibia do not substitute measurements of ultimate load at the radius. The newly proposed multiple section measurements may lead to more robust estimations of bone strength, less affected by subject positioning and

VOI identification, as long as the weakest part of the respective bone for the selected load case is included. Moreover, they provide larger and more flexible common regions for follow-up measurements. *In vivo* bone strength estimated by hFE is an attractive alternative to today's gold standard of μFE models and should be especially encouraged when analyzing multiple sections, due to its much lower repeatability errors and processing time.

CRediT authorship contribution statement

Denis Schenk: Investigation, Data curation, Formal analysis, Writing - original draft, Visualization. **Andrea Mathis:** Investigation. **Kurt Lippuner:** Conceptualization, Resources, Writing - review & editing. **Philippe Zysset:** Conceptualization, Resources, Writing - review & editing, Supervision.

Declaration of competing interest

The authors state no conflict of interest.

Acknowledgment

The authors acknowledge the contributions made by co-workers at the Department of Osteoporosis of the University Hospital in Bern. Additionally, the technical support received from Scanco Medical AG during the study and the advice for the statistical analysis by Prof. Dr. Jasmin Wandel of the Bern University of Applied Sciences is gratefully acknowledged. This work was internally funded by the ARTORG Center for Biomedical Engineering Research and the Department of Osteoporosis of the University Hospital in Bern.

Appendix A. Characteristics of the study population

Table 5
Characteristics of the study population (radius).

		Female	Male
		Mean ± SD	Mean ± SD
Bone mass			
BMC	mgHA	1711 ± 268.5	2689 ± 394.5
Density			
Tot.vBMD	mg HA/cm^3	282.6 ± 59.6	338.8 ± 56.5
Tb.vBMD	mg HA/cm^3	149.7 ± 39.0	193.2 ± 32.0
Ct.vBMD	mg HA/cm^3	890.2 ± 42.2	902.3 ± 24.3
Tb.BV/TV	–	0.2 ± 0.1	0.3 ± 0.1
Geometry			
Tb.Ar	mm^2	246.2 ± 68.1	319.2 ± 71.3
Ct.Ar	mm^2	54.5 ± 16.3	80.4 ± 12.0
Ct.Th	mm	1.02 ± 0.15	1.28 ± 0.19
Microstructure			
Tb.N	1/mm	1.51 ± 0.22	1.65 ± 0.10
Tb.Th	mm	0.22 ± 0.01	0.24 ± 0.02
Tb.Sp	mm	0.64 ± 0.11	0.56 ± 0.05
Ct.Po	%	0.01 ± 0.00	0.01 ± 0.00
Biomechanics			
hFE stiffness	kN/mm	26.7 ± 6.5	46.0 ± 10.3
hFE ultimate load	kN	3.83 ± 1.2	7.0 ± 2.0
μFE stiffness	kN/mm	25.4 ± 5.2	44.1 ± 9.7
μFE ultimate load	kN	2.86 ± 1.0	5.0 ± 1.5

Data are presented as mean ± SD of the mean of repeated measurements per sample.

Table 6
Characteristics of the study population (tibia).

		Female	Male
		Mean ± SD	Mean ± SD
Bone mass			
BMC	mgHA	7285 ± 969.7	11,308 ± 1457.9
Density			
Tot.vBMD	mg HA/cm^3	245.9 ± 34.0	298.2 ± 47.1
Tb.vBMD	mg HA/cm^3	187.5 ± 25.6	236.1 ± 33.7
Ct.vBMD	mg HA/cm^3	807.4 ± 84.6	805.4 ± 39.8
Tb.BV/TV	–	0.27 ± 0.0	0.34 ± 0.1
Geometry			
Tb.Ar	mm^2	875.1 ± 106.6	1122.5 ± 201.0
Ct.Ar	mm^2	92.0 ± 10.7	135.2 ± 24.8
Ct.Th	mm	1.04 ± 0.14	1.39 ± 0.35
Microstructure			
Tb.N	1/mm	1.57 ± 0.21	1.76 ± 0.13
Tb.Th	mm	0.25 ± 0.01	0.27 ± 0.03
Tb.Sp	mm	0.61 ± 0.10	0.53 ± 0.05
Ct.Po	%	0.02 ± 0.01	0.03 ± 0.01
Biomechanics			

(*continued on next page*)

Table 6 (*continued*)

		Female	Male
		Mean ± SD	Mean ± SD
hFE stiffness	kN/mm	43.0 ± 9.3	70.8 ± 13.6
hFE ultimate load	kN	10.7 ± 2.5	18.2 ± 3.9
μFE stiffness	kN/mm	–	–
μFE ultimate load	kN	–	–

Data are presented as mean ± SD of the mean of repeated measurements per sample.

References

[1] B. van Rietbergen, K. Ito, A survey of micro-finite element analysis for clinical assessment of bone strength: the first decade, J. Biomech. 48 (5) (2015) 832–841.

[2] E. J. Samelson et al., "Cortical and trabecular bone microarchitecture as an independent predictor of incident fracture risk in older women and men in the Bone Microarchitecture International Consortium (BoMIC): a prospective study," Lancet Diabetes Endocrinol., vol. 7, no. January, pp. 34–43, 2019.

[4] S. Boutroy, B. Van Rietbergen, E. Sornay-Rendu, F. Munoz, M.L. Bouxsein, P.D. Delmas, Finite element analysis based on in vivo HR-pQCT images of the distal radius is associated with wrist fracture in postmenopausal women, J. Bone Miner. Res. 23 (3) (2007) 392–399.

[5] P. Varga, S. Baumbach, D. Pahr, P.K. Zysset, Validation of an anatomy specific finite element model of Colles' fracture, J. Biomech. 42 (11) (2009) 1726–1731.

[6] P. Varga, D.H. Pahr, S. Baumbach, P.K. Zysset, HR-pQCT based FE analysis of the most distal radius section provides an improved prediction of Colles' fracture load in vitro, Bone 47 (5) (2010) 982–988.

[7] M. Stipsitz, P. K. Zysset, and D. H. Pahr, "Efficient materially nonlinear μ FE solver for simulations of trabecular bone failure," Biomech. Model. Mechanobiol., no. 0123456789, 2019.

[8] M. F. Adams, H. H. Bayraktar, T. M. Keaveny, and P. Papadopoulos, "Ultrascalable Implicit Finite Element Analyses in Solid Mechanics with over a Half a Billion Degrees of Freedom," in Supercomputing, 2004. Proceedings of the ACM/IEEE SC2004 Conference, p. 34.

[9] A.J. Arias-Moreno, et al., Validation of distal radius failure load predictions by homogenized- and micro-finite element analyses based on second-generation high-resolution peripheral quantitative CT images, Osteoporos. Int. 30 (7) (2019) 1433–1443.

[10] W. Pistoia, B. Van Rietbergen, E.-M.M. Lochmüller, C.A. Lill, F. Eckstein, P. Rüegsegger, Estimation of distal radius failure load with micro-finite element analysis models based on three-dimensional peripheral quantitative computed tomography images, Bone 30 (6) (2002) 842–848.

[11] H.S. Hosseini, et al., Fast estimation of Colles' fracture load of the distal section of the radius by homogenized finite element analysis based on HR-pQCT, Bone 97 (2017) 65–75.

[12] S. Boutroy, M.L. Bouxsein, F. Munoz, P.D. Delmas, In vivo assessment of trabecular bone microarchitecture by high-resolution peripheral quantitative computed tomography, J. Clin. Endocrinol. Metab. 90 (12) (2005) 6508–6515.

[13] S. Khosla, et al., Effects of sex and age on bone microstructure at the ultradistal radius: a population-based noninvasive in vivo assessment, J. Bone Miner. Res. 21 (1) (2005) 124–131.

[14] S.F. Baumbach, R. Schmidt, P. Varga, T. Heinz, V. Vécsei, P.K. Zysset, Where is the distal fracture line location of dorsally displaced distal radius fractures? J. Orthop. Res. 29 (4) (2011) 489–494.

[15] T.L. Mueller, M. Stauber, T. Kohler, F. Eckstein, R. Müller, G.H. van Lenthe, Noninvasive bone competence analysis by high-resolution pQCT: an in vitro reproducibility study on structural and mechanical properties at the human radius, Bone 44 (2) (2009) 364–371.

[16] A.K. Stuck, D. Schenk, P. Zysset, L. Bütikofer, A. Mathis, K. Lippuner, Reference Values and Clinical Predictors of Bone Strength for HR-pQCT-Based Distal Radius and Tibia Strength Assessments in Women and Men, Osteoporos. Int, 2020.

[17] S. Bonaretti, S. Majumdar, T.F. Lang, S. Khosla, A.J. Burghardt, The comparability of HR-pQCT bone measurements is improved by scanning anatomically standardized regions, Osteoporos. Int. (2017) 1–14.

[18] W Pistoia, B van Rietbergen, E.M. Lochmüller, C.A. Lill, F Eckstein, P. Rüegsegger, et al., Image-based micro-finite-element modeling for improved distal radius strength diagnosis, J. Clin. Densitom. 7 (2) (2004) 153–160 2004.

[19] P. Varga, E. Dall'Ara, D.H. Pahr, M. Pretterklieber, P.K. Zysset, Validation of an HR-pQCT-based homogenized finite element approach using mechanical testing of ultra-distal radius sections, Biomech. Model. Mechanobiol. 10 (4) (2011) 431–444.

[20] C.C. Glüer, G. Blake, Y. Lu, B.A. Blunt, M. Jergas, H.K. Genant, Accurate assessment of precision errors: how to measure the reproducibility of bone densitometry techniques, Osteoporos. Int. 5 (4) (1995) 262–270.

[21] Y. Pauchard, A.M. Liphardt, H.M. Macdonald, D.A. Hanley, S.K. Boyd, Quality control for bone quality parameters affected by subject motion in high-resolution peripheral quantitative computed tomography, Bone 50 (6) (2012) 1304–1310.

[22] R. Ellouz, R. Chapurlat, B. van Rietbergen, P. Christen, J.B. Pialat, S. Boutroy, Challenges in longitudinal measurements with HR-pQCT: evaluation of a 3D registration method to improve bone microarchitecture and strength measurement reproducibility, Bone 63 (2014) 147–157.

[23] J.A. MacNeil, S.K. Boyd, Improved reproducibility of high-resolution peripheral quantitative computed tomography for measurement of bone quality, Med. Eng. Phys. 30 (6) (2008) 792–799.

[25] M.A. Paggiosi, R. Eastell, J.S. Walsh, Precision of high-resolution peripheral quantitative computed tomography measurement variables: influence of gender, examination site, and age, Calcif. Tissue Int. 94 (2) (2014) 191–201.

[26] A.J. Burghardt, et al., Multicenter precision of cortical and trabecular bone quality measures assessed by high-resolution peripheral quantitative computed tomography, J. Bone Miner. Res. 28 (3) (2013) 524–536.

[27] J.B.B. Pialat, A.J. Burghardt, M. Sode, T.M. Link, S. Majumdar, Visual grading of motion induced image degradation in high resolution peripheral computed tomography: impact of image quality on measures of bone density and micro-architecture, Bone 50 (1) (2012) 111–118.

[28] M.L. Bouxsein, S.K. Boyd, B.A. Christiansen, R.E. Guldberg, K.J. Jepsen, R. Müller, Guidelines for assessment of bone microstructure in rodents using micro-computed tomography, J. Bone Miner. Res. 25 (7) (2010) 1468–1486.

[29] T.P. Harrigan, R.W. Mann, Characterization of microstructural anisotropy in cancellous bone using a second rank tensor, J. Mater. Sci. 19 (1985) 761–767.

[30] H.S. Hosseini, M. Horak, P.K. Zysset, M. Jirasek, An over-nonlocal implicit gradient-enhanced damage-plastic model for trabecular bone under large compressive strains, Int. J. Numer. Methods Biomed. Eng. 26 (1) (2015) 807–827.

[31] K. Chiba, et al., Precision of second-generation high-resolution peripheral quantitative computed tomography: intra- and intertester reproducibilities and factors involved in the reproducibility of cortical porosity, J. Clin. Densitom. 21 (2) (2018) 295–302.

[32] N. Mikolajewicz, et al., HR-pQCT Measures of Bone Microarchitecture Predict Fracture: Systematic Review and Meta-analysis, J. Bone Miner. Res. (2019) 1–14, https://doi.org/10.1002/jbmr.3901.

[33] S.L. Manske, E.M. Davison, L.A. Burt, D.A. Raymond, S.K. Boyd, The estimation of second-generation HR-pQCT from first-generation HR-pQCT using in vivo cross-calibration, J. Bone Miner. Res. 32 (7) (2017) 1514–1524.

[34] M. Sode, A.J. Burghardt, J.B. Pialat, T.M. Link, S. Majumdar, Quantitative characterization of subject motion in HR-pQCT images of the distal radius and tibia, Bone 48 (6) (2011) 1291–1297.

[35] S. Caksa, A. Yuan, S. E. Rudolph, E. W. Yu, K. L. Popp, and M. L. Bouxsein, "Influence of soft tissue on bone density and microarchitecture measurements by high-resolution peripheral quantitative computed tomography," Bone, vol. 124, no. March, pp. 47–52, 2019.

[36] C.E. Kawalilak, S.A. Kontulainen, M.A. Amini, J.L. Lanovaz, W.P. Olszynski, J.D. Johnston, In vivo precision of three HR-pQCT-derived finite element models of the distal radius and tibia in postmenopausal women, BMC Musculoskelet. Disord. 17 (1) (2016) 1–11.

[37] S. Bonaretti, et al., Operator variability in scan positioning is a major component of HR-pQCT precision error and is reduced by standardized training, Osteoporos. Int. 28 (1) (2017) 245–257.

[38] C.E. Kawalilak, J.D. Johnston, W.P. Olszynski, D.A. Leswick, S.A. Kontulainen, Comparison of short-term in vivo precision of bone density and microarchitecture at the distal radius and tibia between postmenopausal women and young adults, J. Clin. Densitom. 17 (4) (2014) 510–517.